UQOP 2019
Uncertainty Quantification & Optimization Conference
18–20 March, Paris, France

Near-optimal smoothing in derivative-free stochastic optimization

Friedrich Menhorn*, Gianluca Geraci◇, Michael S. Eldred◇,
Hans-Joachim Bungartz*, Youssef M. Marzouk†

*Technical University of Munich, †Massachusetts Institute of Technology,
◇Sandia National Laboratories
menhorn@in.tum.de

Abstract

SNOWPAC (Stochastic Nonlinear Optimization With Path-Augmented Constraints) solves optimization under uncertainty (OUU) problems where the underlying problem is only available as black box, using a derivative-free trust region approach. Monte Carlo estimators are used to evaluate the objective and constraint functions of the OUU formulation. Fully linear models for noisy functions and Gaussian process surrogate models smooth these noisy function evaluations. We use these ingredients to develop near-optimal noise correction schemes for SNOWPAC. We also describe the use of multifidelity Monte Carlo estimators, and their error estimates, to reduce the overall computation cost of optimization. We report results on benchmark problems and demonstrate the entire framework in the computationally challenging design of a scramjet combustor.

I. Introduction

The stochastic optimization method SNOWPAC [Augustin & Marzouk, 2017] extends the deterministic derivative-free trust region method NOWPAC [Augustin & Marzouk, 2014], which proposes a new way of handling nonlinear constraints via an "inner boundary path" that guarantees feasibility. In particular, NOWPAC employs an additive function to convexify nonlinear constraints that are known only through pointwise evaluations, thus yielding feasible trial steps and global convergence to a local first-order minimum.

SNOWPAC tackles OUU problems formulated as follows:

$$\begin{aligned} &\min \mathcal{R}^f_\pi(\mathbf{x},\theta), \\ \text{s.t.} \quad &\mathcal{R}^{c_i}_\pi(\mathbf{x},\theta) \le 0,\ i = 1,\dots,r. \end{aligned} \tag{1}$$

with design parameters $\mathbf{x} \in \mathbb{R}^d$ and uncertain parameters $\theta \sim \pi$, where π is a probability distribution over $\Theta \subseteq \mathbb{R}^k$. Here $\mathcal{R}^f_\pi$ and $\mathcal{R}^c_\pi$ are measures of robustness or risk derived from the objective function $f : \mathbb{R}^d \times \Theta \to \mathbb{R}$ and nonlinear inequality constraints $c_i : \mathbb{R}^d \times \Theta \to \mathbb{R}$, $i = 1 \dots r$, respectively. An example robustness measure is a linear combination of mean $\mathbb{E}_\pi[\,\cdot\,]$

and variance $\mathbb{V}_\pi[\cdot]$, which accounts for the average and spread of possible realizations. Other options include event probabilities (yielding chance constraints) or conditional value-at-risk.

II. Methods

To evaluate the measures $\mathcal{R}^b_\pi$ given in (1), we use Monte Carlo estimators R^b (where b is f or c_i). Doing so introduces an error that depends on the number of samples N, i.e., $\mathcal{R}^b = R^b + \varepsilon^b_N$ for some random variable ε^b_N. A first step to mitigating the impact of this error is to employ the minimum Frobenius norm surrogate models m^b_k introduced in [Kannan and Wild, 2012], in a neighborhood (called the trust region) of size ρ_k around the current design $\mathbf{x}_k$. The maximum magnitude of the noise yields a lower bound on the trust region radius, $\rho_k \geq \lambda\sqrt{\varepsilon_{\max,k}} = \max_i \lambda\sqrt{\varepsilon_N(\mathbf{x}_{k,i})}$, for $\lambda \in (0,\infty)$ and evaluation points $\{\mathbf{x}_{k,i}\}_{i=1}^M$ at each optimization step k.

To allow the trust region radius to shrink, SNOWPAC uses Gaussian process (GP) surrogates to gradually *smooth* the noisy function evaluations $R^b_{k,i} := R^b(\mathbf{x}_{k,i})$, replacing them with

$$\widetilde{R}^b_{k,i} = \gamma_{k,i}\mu^{GP}_{k,i} + (1-\gamma_{k,i})R^b_{k,i}, \tag{2}$$

where $\mu^{GP}_{k,i} := \mu^{GP}(\mathbf{x}_{k,i}; \{R^b_i\}_{i=1}^M)$ denotes mean of a GP trained on all available evaluations $\{R^b_i\}_{i=1}^M$, and $\gamma_{k,i} \in [0,1]$ is a mixing weight. We show that an optimal weight can be found by calculating

$$\gamma_{k,i} = \frac{\mathbb{V}[R^b_{k,i}] - \mathbb{C}\mathrm{ov}[\mu^{GP}_{k,i}, R^b_{k,i}]}{(\mathbb{E}[\mu^{GP}_{k,i}] - \mathcal{R}^b_{k,i})^2 + \mathbb{V}[\mu^{GP}_{k,i} - R^b_{k,i}]}. \tag{3}$$

This choice minimizes the magnitude of the noise $\varepsilon_{k,i}$, i.e., the mean squared error (MSE) of our estimator

$$\mathrm{MSE}(\widetilde{R}^b_{k,i}) = [\gamma_{k,i}(\mathbb{E}[\mu^{GP}_{k,i}] - \mathcal{R}^b_{k,i})]^2 + \mathbb{V}[\gamma_{k,i}\mu^{GP}_{k,i} + (1-\gamma_{k,i})R^b_{k,i}]. \tag{4}$$

We present approximations of (3) that provide near-optimal reduction of the MSE in each step.

III. Results

We validate our smoothing approach using OUU problems derived from the CUTEst benchmark suite showing improved results compared to other optimization methods. Additionally, we demonstrate the use of multilevel Monte Carlo estimators of the robustness measures, and their recently developed error estimators (see [Menhorn et al., 2018]), for an artificial turbulent flow scenario and a realistic scramjet design optimization problem.

References

[Augustin & Marzouk, 2014] Augustin, F. and Marzouk, Y. M. (2014), A Path-Augmented Constraint Handling Approach for Nonlinear Derivative-Free Optimization, *arXiv*, 1403.1931v3

[Augustin & Marzouk, 2017] Augustin, F. and Marzouk, Y. M. (2017), A Trust-Region Method for Derivative-Free Nonlinear Constrained Stochastic Optimization, *arXiv*, 1703.04156

[Kannan and Wild, 2012] Kannan, A and Wild, S.M. (2012), Obtaining Quadratic Models of Noisy Functions, *Technical Report ANL/MCS-P1975-1111*, Argonne National Laboratory, USA

[Menhorn et al., 2018] Menhorn, F., Geraci G., Eldred M.S. and Marzouk Y.M. (2018), Multifidelity OUU for a Scramjet-Inspired Problem, *ECCM-ECFD 2018*, Glasgow, UK